Recombinant bacterial strains for the production of natural nucleosides and modified analogues thereof.

The present invention relates to novel genetically modified bacterial strains capable of expressing polypeptides having the enzyme activity of the enzymes UdP and PNP; the strains in question can be used to catalyse transglycosylation reactions between a donor nucleoside and an acceptor base.

Natural nucleosides or the modified analogues thereof have important applications, both directly and as intermediates, in the field of drugs having an anti-viral and anti-tumour action, as well as in the preparation of oligonucleotides for therapeutic and diagnostic use.

Nucleosides can be prepared using methods of chemical synthesis which normally require a large number of steps processes for the protection and deprotection of labile groups and the use of reagents and operating conditions which, on an industrial level, may be both difficult to apply and economically disadvantageous. In addition, those reactions do not generally have high overall yields owing also to the formation of mixtures of stereo- and regio-isomers from which the compound of interest has to be separated.

An alternative approach to the preparation of nucleosides and modified analogues thereof is based on interconversion between a sugar-donating nucleoside and an acceptor base by means of enzymes which catalyse the general reversible reactions (Hutchinson, Trends Biotechnol. 8, 348-353, 1990) given below in scheme 1:

Scheme 1

pyrimidine-β-(deoxy)ribonucleoside + Pi

Reaction 1 ↑↓

pyrimidine base + α-(deoxy)ribose-1-phosphate

purine base + α-(deoxy)ribose-1-phosphate

Reaction 2 ↑↓

purine-β-(deoxy)ribonucleoside + Pi

where Pi = organic phosphate.

Reaction 1 is catalysed by the enzyme uridine phosphorylase or UdP (E.C.2.4.2.3.) while reaction 2 is catalysed by the enzyme purine nucleoside phosphorylase or PNP (E.C.2.4.2.1.).

The UdP and PNP enzymes can be used individually to catalyse transglycosylation reactions between a donor pyrimidine nucleoside and an acceptor pyrimidine base or between a donor purine nucleoside and an acceptor purine base, respectively. In addition, when the two enzymes are used in combination, it is possible to transfer the sugar from a donor pyrimidine nucleoside to a purine or pyrimidine acceptor base as well as from a donor purine nucleoside to a pyrimidine or purine acceptor base, depending on the starting materials used. In each case the phosphorolysis reactions involve a configuration change at position 1 of the sugar to give an α-sugar-1-phosphate which constitutes the intermediate substrate of the transglycosylation reactions and which is subsequently transferred to the acceptor base, with restoration of the original β configuration.

Those enzyme reactions can advantageously be carried out starting from a mixture of a donor nucleoside and an acceptor base in the simultaneous presence of the two enzymes and without isolating the intermediate sugar phosphate or in two steps comprising phosphorolysis with formation of the intermediate sugar phosphate, its isolation and subsequent condensation with the acceptor base.

With regard to chemical synthesis, an important advantage of transglycosylation reactions catalysed by phosphorylases is the maintenance of stereo-selectivity and regio-selectivity, as a result of which the end product retains the β configuration of natural nucleosides.

The UdP and PNP enzymes which participate physiologically in the catabolism and interconversion reactions of nucleosides are the product, respectively, of the *udp* and *deoD* genes, occurring widely in nature, and have been identified and studied in both prokaryotic and eukaryotic organisms (Parks and Agarwal, Enzymes 7, 3rd ed., 483-514, Academic Press, New York; Munch-Petersen, Metabolism of nucleotides, nucleosides and nucleobases in micro-organisms, Academic Press, London, 1982).

From the point of view of use as catalysts for the synthesis of nucleosides and modified analogues thereof, the enzymes of prokaryotic organisms are generally preferred because they have a lower substrate specificity and they can catalyse transglycosylation reactions

starting also from donor nucleosides containing modified sugars and from acceptor bases comprising both purine or pyrimidine structures and various nitrogen-containing heterocyclic systems (Stoeckler *et al.*, Biochemistry 19, 102-107, 1980; Browska *et al.*, Z. Naturforsch., 45, 59-70, 1990).

The transglycosylation reactions can be carried out using purified or partially purified enzyme preparations (Krenitsky *et al.*, Biochemistry 20, 3615-3621, 1981; EP-002192) or, alternatively, using the whole bacterial cells of microorganisms selected because they contain the necessary enzymes (Utagawa *et al.*, Agric.Biol.Chem. 49, 3239-3246,1985) or whole cells cultivated in the presence of inducers of the production of those enzymes (Doskocil *et al.*, Collect. Czech. Chem. Commun. 42, 370-383, 1977).

For biocatalysis reactions carried out at a preparative level, the use of whole cells both obviates the need to extract and purify the enzymes and enables the cells to be recovered easily at the end of the reaction, for example by centrifugation or ultrafiltration, and to be re-used for other, subsequent, reaction cycles; alternatively, it is possible to use the UdP and PNP enzymes extracted from the cells in the form of a crude or purified soluble cell fraction. Both UdP and PNP are enzymes characterised by good thermal stability which enables the transglycosylation reactions to be carried out at temperatures of up to approximately 60°C without significant activity losses and enables the recovered enzyme preparations to be re-used. Approaches have also been described where the recycling of cells used as catalysts was carried out by micro-encapsulation in both hydrophilic gels (Votruba *et al.*, Collect.Czech.Chem. Commun. 59, 2303-2330, 1994) and hydrophobic gels (Yokozeki *et al.*, Eur.J.Appl.Microbiol. Biotechnol., 14, 225-231, 1982).

The main limitations of the methods known hitherto for the preparation of natural nucleosides and modified analogues thereof by transglycosylation reactions using bacterial cells reside in the low enzyme concentration obtainable, even after induction, and in the impossibility of using optimised amounts of the two enzyme activities required to catalyse the transfer of the sugar from a donor nucleoside to an acceptor base.

Both in the case of selection of wild-type bacterial strains and in the case of cultivation of strains under induction conditions, cells are obtained that contain levels of UdP and PNP which are generally not higher than 10 times the base levels (F.Ling *et al.*, Process Biochem. 29,355-361,1994) and which are in non-predeterminable ratios. Furthermore,

because one of the two enzymes (generally PNP) is present in the induced cells in lower amounts, it is usually necessary to use an excess of cells such as to ensure the presence of the limiting enzyme at levels compatible with acceptable overall kinetics of the interconversion reaction. From an operating point of view, this means that a significant portion of the reaction mixture is constituted by the cell suspension, with consequent restriction of the volume that can be used to solubilise the substrates and, finally, with a lower volumetric yield of end product.

The present invention therefore relates to the construction of genetically modified bacterial strains capable of solving the problems described above and, in particular, of catalysing transglycosylation reactions between a donor nucleoside and an acceptor base with high yields which are foreseeable and, above all, reproducible on an industrial scale and with particularly rapid enzyme kinetics.

The literature has described the cloning and expression of some recombinant phosphorylases, such as, for example, human UdP (Watanabe and Uchida, Biochem.Biophys.Res.Commun. 216, 265-272, 1996), murine UdP (Watanabe *et al.*, J.Biol.Chem. 270, 12191-12196, 1995), of *Escherichia coli* (Mikhailov *et al.*, Biochem.Internat. 26, 607-615, 1992) and human PNP (Erion *et al.*, Biochemistry 36, 11725-11734, 1997), of the thermophilic micro-organism *Bacillus stearothermophilus* (Hamamoto *et al.*, Biosci.Biotech.Biochem. 61, 272-275, 1997; Hamamoto *et al.*, Biosci. Biotech. Biochem. 61, 276-280, 1997) in addition to UdP and PNP from *Klebsiella* sp (Takehara *et al.*, Biosci.Biotech.Biochem. 59, 1987-1990, 1995). In particular, Japanese patent application JP-06-253854 describes the expression in *E.coli* of bacterial plasmids containing the gene sequences of the enzymes purine and/or pyrimidine nucleoside phosphorylase derived solely from thermophilic bacteria, that is bacteria having optimum growth at temperatures of from 50 to 60°C, such as, for example, *Bacillus stearotermophilus*.

Novel genetically modified bacterial strains that contain the genes coding for polypeptides having the enzyme activity of the enzymes uridine phosphorylase (UdP) and/or purine nucleoside phosphorylase (PNP), both separately and together, have now been found and constitute part of the subject matter of the present invention. The cultivation of these novel strains enables both high levels of biomass and high levels of expression of the recombinant

enzymes to be obtained; the novel strains according to the present invention can also be used either directly or after extraction of the soluble cell fraction as catalysts for the production of natural nucleosides and modified analogues thereof with substantial improvements in the process in comparison with the prior art.

In contrast to what has been described in JP-06-253854, the plasmid vectors according to the present invention can be obtained by cloning both separately and simultaneously the *udp* and *deoD* genes of mesophilic bacteria, that is bacteria having optimum growth at temperatures of from 30 to 37°C such as, for example, *E.coli*. To be more precise, the gene sequences preferably used for the purposes of the present invention are the *E.coli* sequences that encode the *udp* and *deoD* genes and that are deposited in the EMBL data bank with the accession numbers X15689 (*udp*) and M60917 (*deoD*); however, it is also possible to use other widely available sequences, such as, for example, AC CG01747 (*udp*) and AC CG00327 (*deoD*).

The expression plasmid vectors which may be used for the purposes of the invention and which form part of the subject matter thereof are therefore characterised in that they comprise:

a) at least one gene sequence of a mesophilic bacterium coding for a polypeptide having enzyme UdP and/or enzyme PNP activity; and

b) at least one gene sequence coding for antibiotic resistance.

The at least one sequence coding for antibiotic resistance is preferably a sequence coding for tetracycline, kanamycin and/or ampicillin resistance. The plasmid vectors of the present invention can be obtained by cloning either the sequence coding for *udp* and/or the sequence coding for *deoD* or, optionally, the sequence coding for tetracycline and/or kanamycin resistance into the plasmid pUC18 (Yanish and Perron, Gene 33, 103-119, 1985; EMBL accession number L08752) which already contains the ampicillin resistance gene.

The relative position of the sequences coding for *udp* and *deoD* is not, however, relevant for the purposes of the invention: that is to say, the sequence coding for *udp* can be positioned either downstream or upstream of the sequence coding for *deoD*. Furthermore, and as it will be appreciated from the Examples which follow, the gene sequences coding

for *udp* and *deoD* may also be fused together so to express novel fusion proteins wherein the enzymes UdP and PNP are either covalently bonded together (UdP-PNP) or, alternatively, the novel fusion protein may have the formula UdP-(L)-PNP wherein L is a polipeptide linker of more than one aminoacidic unit. In these novel fusion proteins, the relative position of the two components is not however relevant for the purposes of the invention: that is to say, the PNP component can be either at the NH_2-terminal or at the COOH-terminal position of the fused proteins. The novel proteins thus obtainable, which are a further object of the present invention, are characterized by possessing a bifunctional activity as they are able to perform both the activity of the enzyme UdP and that of the enzyme PNP.

An additional object of the present invention is then represented by the method for producing the above mentioned fusion proteins, said method comprising:

(a) producing a plasmid expression vector as above indicated;

(b) transforming a host bacteria cell with said expression vector; and

(c) isolating and purifying the fusion protein from the transformed bacteria cell.

The methods for transforming a host bacteria cell with an expression vector and for isolating and purifying the expressed peptide are well known to any skilled in this art and are for example disclosed in Swartz JR, *Escherichia coli* recombinant DNA technology, and in Neidahrt FC et al. (edts), *Escherichia coli* and *Salmonella typhimurium*: Cellular and molecular biology, 2nd edition, pp 1693-1711, ASM, Washington, herein incorporated as a reference.

The hosts preferably used for the expression of the recombinant enzymes according to the present invention are bacterial cells of *Escherichia coli*; the strains K12 (preferably DH5α or MG1655) and/or the B strains are of particular interest. Alternatively, however, it is possible to use cells of other prokaryotic micro-organisms which are acceptable for industrial use because they are not dangerous to operators and the environment and they can be readily cultivated to obtain high levels of biomass.

As will also be seen from the Examples, the presence of a bacterial promoter, and in particular of the *lac* promoter, is not an essential element for the purposes of the present invention because it has been found that cell growth and the expression of polypeptides do

not depend on the presence of an inducer (IPTG). For ease of performance, the gene sequence encoding a polypeptide having enzyme UdP activity and/or enzyme PNP activity is cloned into the plasmid pUC18 in the reading frame relative to the *lac* promoter.

Finally, the sequence coding for tetracycline resistance is preferably the Tet gene of pBR322; the sequence coding for kanamycin resistance is the kan gene of pET29c.

Thus, in accordance with well-known methods which will become clear from the Examples, the following plasmids, which are represented in figures 1, 3 and 4, were constructed:

- pGM679: *udp* gene cloned into plasmid pUC18 (SEQ ID NO 1). In the sequence numbering, coordinate 1 of pGM679 coincides with that of the pUC18 vector sequence; from nucleotide 1 to 242: pUC18 sequence; from 243 to 1021: *E.coli udp* gene sequence; from 1022 to 3444: pUC18 sequence.
- pGM708: *udp* gene cloned into plasmid pUC18 together with the tetracycline resistance gene (SEQ ID NO 2). In the sequence numbering, coordinate 1 of pGM708 coincides with that of the pUC18 vector sequence; from nucleotide 1 to 242: pUC18 sequence; from 243 to 1021: *E.coli udp* gene sequence; from 1022 to 1039: pUC18 sequence; from 1040 to 1482: pHP45Ω sequence; from 1483 to 2883: pBR322 Tet gene sequence; from 2884 to 3151: pHP45Ω sequence; from 3152 to 5556: pUC18 sequence.
- pGM678: *deoD* gene cloned into plasmid pUC18 (SEQ ID NO 3). In the sequence numbering, coordinate 1 of pGM678 coincides with that of the pUC18 vector sequence; from nucleotide 1 to 230: pUC18 sequence; from 231 to 960: *E.coli deoD* gene sequence; from 961 to 3383: pUC18 sequence.
- pGM707: *deoD* gene cloned into plasmid pUC18 together with the tetracycline resistance gene (SEQ ID NO 4). In the sequence numbering, coordinate 1 of pGM707 coincides with that of the pUC18 vector sequence; from nucleotide 1 to 230: pUC18 sequence; from 231 to 960: *E.coli deoD* gene sequence; from 961 to 978: pUC18 sequence; from 979 to 1422: pHP45Ω sequence; from 1423 to 2822: pBR322 Tet gene sequence; from 2823 to 3090: pHP45Ω sequence; from 3091 to 5495: pUC18 sequence.

- pGM712: *udp* and *deoD* genes cloned into plasmid pUC18 (SEQ ID NO 5). In the sequence numbering, coordinate 1 of pGM712 coincides with that of the pUC18 vector sequence; from nucleotide 1 to 242: pUC18 sequence; from 243 to 1021: *E.coli udp* gene sequence; from 1022 to 1025: pUC18 sequence; from 1026 to 1036: pBAD24 sequence; from 1037 to 1766: *E.coli deoD* gene sequence; from 1767 to 1792: pBAD24 sequence; from 1793 to 4189: pUC18 sequence.

- pGM716: *udp* and *deoD* genes cloned into plasmid pUC18 together with the tetracycline resistance gene (SEQ ID NO 6). In the sequence numbering, coordinate 1 of pGM716 coincides with that of the pUC18 vector sequence; from nucleotide 1 to 242: pUC18 sequence; from 243 to 1021: *E.coli udp* gene sequence; from 1022 to 1025: pUC18 sequence; from 1026 to 1036: pBAD24 sequence; from 1037 to 1766: *E.coli deoD* gene sequence; from 1767 to 1792: pBAD24 sequence; from 1793 to 1794: pUC18 sequence; from 1795 to 2228: pHP45Ω sequence; from 2229 to 3628: pBR322 Tet gene sequence; from 3629 to 3896: pHP45Ω sequence; from 3897 to 6301: pUC18 sequence.

- pGM709: gene *deoD* cloned in pBAD24 (SEQ ID NO 7). In the sequence numbering, coordinate 1 of pGM709 coincides with that of the pBAD24 vector sequence; from nucleotide 1 to 1311: pBAD24 sequence; from 1312 to 2042: sequence corresponding to 230-960 of pGM678; from 2043 to 5241: pBAD24 sequence.

- pGM769: pGM716 with deletion of HpaI fragment (SEQ ID NO 8). In the sequence numbering, coordinate 1 of pGM769 coincides with that of pGM716 sequence; from nucleotide 1 to 914: pGM716 sequence; from nucleotide 915 to 5822: sequence corresponding to 1394-6301 of pGM716.

- pGM771: genes *udp* and *deoD* cloned in pUC18 so to create a fusion between the two proteins; the plasmid also bears the tetracycline resistance gene (SEQ ID NO 9). In the sequence numbering, coordinate 1 of pGM771 coincides with that of pGM716 sequence; from nucleotide 1 to 1011: pGM716 sequence; from nucleotide 1012 to 6269: sequence corresponding to 1044-6301 of pGM716.

- pGM795: genes *udp* and *deoD* cloned in pUC18 so to create a fusion between the two proteins bonded to each other via an aminoacidic linker; the plasmid also bears the

tetracycline resistance gene (SEQ ID NO 10). In the sequence numbering, coordinate 1 of pGM795 coincides with that of pGM716 sequence, from nucleotide 1 to 1011: pGM771 sequence; from 1012 to 1041: linker sequence; from 1042 to 6299: sequence corresponding to 1044-6301 of pGM716.

- pGM746: cloning vector derived from pUC18 (SEQ ID NO 11). In the sequence numbering, coordinate 1 of pGM746 coincides with that of the pUC18 vector sequence; from nucleotide 1 to 54: pUC18 sequence; from 55 to 109: pUC18 polylinker sequence; from 110 to 2297 pUC18 sequence.

- pGM747: *deoD* gene cloned into pGM746 without upstream promoter (SEQ ID NO 12). In the sequence numbering, coordinate 1 of pGM747 coincides with that of pGM746; from nucleotide 1 to 79: pGM746 sequence; from 80 to 837:sequence corresponding to 1301-2058 of pGM709; from 838 to 3031: pGM746 sequence.

- pGM751: *deoD* gene cloned downstream promoter *ptac* (SEQ ID NO 13). In the sequence numbering, coordinate 1 of pGM751 coincides with that of pGM747; from nucleotide 1 to 72: pGM747 sequence; from 73 to 171: *ptac* sequence from pGZ119; from 172 to 3128: pGM747 sequence.

- pGM800: genes *udp* and *deoD* cloned downstream *ptac* promoter into a vector derived from pUC18 (SEQ ID NO 14). In the sequence numbering, coordinate 1 of pGM800 coincides with that of pGM751; from nucleotide 1 to 923: pGM751 sequence; from 924 to 1741: *udp* sequence corresponding to 203-1020 of pGM679; from 1742 to 3934: pGM751 sequence.

- pGM807: genes *udp* and *deoD* cloned downstream *ptac* promoter into a vector containing the tetracycline resistance gene (SEQ ID NO 15). In the sequence numbering, coordinate 1 of pGM807 coincides with that of pGM800; from nucleotide 1 to 1742: pGM800 sequence; from 1743 to 3855: Tc sequence from pHP45α; from 3856 to 6046: pGM800 sequence.

The recombinant strains so obtained express polypeptides having enzyme UdP and PNP activity in large amounts, minimising any compatibility and/or solubility problems which can be caused by the presence of heterologous proteins.

In particular, the bacterial strains called DH5α/pGM678, MG1655/pGM678, DH5α/pGM707 and MG1655/pGM707 which overexpress the enzyme PNP; the strains DH5α/pGM679, MG1655/pMG679, DH5α/pGM708 and MG1655/pGM708 which overexpress the enzyme UdP; the strains DH5α/pGM712, DH5α/pGM716, MG1655/pGM716, DH5α/pGM800 and DH5α/pGM807 which overexpress the enzymes PNP and UdP simultaneously in the same cell; and the strains DH5α/pGM771, MG1655/pGM771, DH5α/pGM795, MG1655/pGM795, which overexpress the bifunctional fusion proteins UdP-(L)-PNP, were constructed. The efficiency of these novel strains, both as producers of the enzymes PNP and UdP and as biocatalysts for the preparation of nucleosides by bioconversion reactions, was compared with a preparation of *Enterobacter aerogenes* cells cultivated in the presence of inducers because that micro-organism, according to the data available in the literature, has hitherto been regarded as one of the best for catalysing transglycosylation reactions (Utagawa *et al.*, Agric.Biol.Chem. 49, 1053-1058, 1985; Utagawa *et al.*, Agric.Biol.Chem. 49, 2711-2717, 1985). The present invention relates also to the use of the novel recombinant strains in the production of polypeptides having enzyme UdP activity and/or enzyme PNP activity and/or as catalysts of transglycosylation reactions between a donor nucleoside and an acceptor base.

The enzyme activity of the recombinant strains was determined by incubating directly the cell suspension, or cell extracts obtained by mechanical and/or enzymatic lysis, in phosphate buffer with a pyrimidine nucleoside (for example uridine) to test for UdP activity or with a purine nucleoside (for example inosine) to test for PNP activity and by determining the formation of the pyrimidine base (uracil) or purine base (hypoxanthine), respectively, by reverse phase high pressure liquid chromatography (RP-HPLC), as indicated in Example 7.

Applying that test, the enzyme activities of UdP and PNP were measured in the recombinant bacterial strains to which the present invention relates and in the comparison *E.aerogenes* strain, to give the results indicated in Tables 1 and 2, which show that the recombinant strains of the present invention have enzyme activities up to approximately 10-30 times higher than that of the comparison strain cultivated under induction conditions and up to approximately 120-1000 times higher than that of the non-transformed *E.coli* host strains.

Table 1. Comparison of the enzyme activities of uridine phosphorylase (UdP) and purine nucleoside phosphorylase (PNP) in recombinant *E.coli* strains and in the comparison *E.aerogenes* strain.

Novel bacterial strains according to the invention	UdP activity units/g of wet cells	PNP activity units/g of wet cells
wild-type MG1655	4.5 ± 0.2	3.5 ± 0.2
MG1655/pGM707	7.5 ± 0.1	2400.0 ± 50.0
MG1655/pGM708	1550.0 ± 60.0	6.5 ± 0.5
MG1655/pGM716	5400.0 ± 450.0	850.0 ± 30.0
Comparison strain		
Non-induced *E.aerogenes* ATCC 13048	3.7 ± 0.2	3.0 ± 0.2
Induced *E.aerogenes* ATCC 13048	168.3 ± 2.9	19.0 ± 2.2

Table 2. Comparison of the enzyme activities of uridine phosphorylase (UdP) and purine nucleoside phosphorylase (PNP) assayed into the cell extracts of the recombinant *E.coli* strains MG1655 and DH5α, in the corresponding wild-type strains and in the non-induced and induced comparison *E.aerogenes* strains.

Novel bacterial strains according to the invention	UdP activity units/g of wet cells	PNP activity units/g of wet cells
non-transformed MG1655	9 ± 0.4	5 ± 0.3
MG1655/pGM707	15 ± 0.2	996 ± 29
MG1655/pGM708	3100 ± 120	10 ± 0.7
MG1655/pGM716	6000 ± 160	643 ± 11

non-transformed DH5α	10 ± 1.0	3 ± 0.2
DH5α/pGM707	14 ± 0.2	1000 ± 20
DH5α/pGM708	10400 ± 750	4 ± 0.6
DH5α/pGM716	6200 ± 150	600 ± 10
E. aerogenes ATCC 13048	7.4 ± 0.4	4.5 ± 0.3
Induced *E. aerogenes* ATCC 13048	335 ± 5	29 ± 3.3

The surprisingly high level of enzyme activity of these novel recombinant strains is confirmed by an indirect comparison with the strains described in JP-06-253854: the strains considered in the present invention permit enzyme activities from 340 to 1040 times (as regards the activity of UdP) and from 120 to 200 times (as regards the activity of PNP) higher than the enzyme activities of the non-transformed wild-type strains; the strains described in JP-06-253854, on the other hand, have an enzyme activity in *E.coli* 150 and 91 times higher, respectively, than that of the corresponding wild-type strain. It is also noteworthy that the enzyme activity of the strains of the present invention was determined at 30°C while that of the strains of JP-06-253854 was established while operating at 70°C, or at a temperature which permits markedly higher kinetics.

This high level of enzyme activity is also confirmed by the overexpression of the enzymes UdP and PNP which can be demonstrated both by electrophoretic analysis (Figure 5) and by quantitative determination by RP-HPLC analysis which demonstrated levels of specific expression of from 55 to 120 milligrams of UdP/gram of wet cell paste and/or from 15 to 65 milligrams of PNP/gram of wet cell paste, as indicated in the example of Table 3.

Table 3

Quantitative determination of UdP and PNP expression levels by RP-HPLC analysis.

Bacterial strains of the present invention	mg UdP/g wet cell paste	mg PNP/g wet cell paste
MG1655/pGM707	-	60
MG1655/pGM716	55	15

DH5α/pGM707	-	65
DH5α/pGM708	120	-
DH5α/pGM716	60	15

The whole cells of the recombinant strains described in the present invention, or their crude or purified extracts, can advantageously be used as biocatalysts for the preparation of natural nucleosides and modified analogues thereof starting from a sugar-donating nucleoside and from an acceptor base by means of bioconversion reactions which require the presence of only one type of phosphorylase (UdP or PNP) or the simultaneous presence of UdP and PNP according to the following general schemes:

a) pyrimidine nucleoside P1 + pyrimidine base P2 → pyrimidine nucleoside P2 + pyrimidine base P1, in the presence of recombinant cells that overexpress UdP;

b) purine nucleoside P1 + purine base P2 → purine nucleoside P2 + purine base P1, in the presence of recombinant cells that overexpress PNP;

c) pyrimidine nucleoside + purine base → purine nucleoside + pyrimidine base, in the presence of a mixture of recombinant cells that overexpress UdP and PNP separately or of cells of a single recombinant strain that co-expresses UdP and PNP;

d) purine nucleoside + pyrimidine base → pyrimidine nucleoside + pyrimidine base, in the presence of a mixture of recombinant cells that overexpress UdP and PNP separately or of cells of a single recombinant strain that co-express UdP and PNP.

According to the information given in the literature, in the bioconversion reactions catalysed by UdP and PNP, there come into consideration as donor nucleosides both natural or modified nucleosides containing D-ribose and 2'-deoxyribose, and nucleosides containing the ribose group modified in the 2', 3' and/or 5' positions and, in particular, nucleosides in which the sugar is constituted by β-D-arabinose, α-L-xylose, 3'-deoxyribose, 3',5'-dideoxyribose, 2',3'-dideoxyribose, 5'-deoxyribose, 2',5'-dideoxyribose, 2'-amino-2'-deoxyribose, 3'-amino-3'-deoxyribose, 2'-fluoro-2'-deoxyribose. The acceptor bases which can be used in the bioconversion reactions catalysed by UdP and PNP are natural or substituted pyrimidine and purine bases, in particular purine bases substituted in the 1, 2 and/or 6 positions, pyrimidine bases

substituted in the 3 and/or 5 positions and also other heterocyclic systems containing one or more nitrogen atoms, such as, for example, purine, 2-azapurine, 8-azapurine and substituted analogues thereof, 1-deazapurine (imidazopyridine), 3-deazapurine, 7-deazapurine and substituted analogues thereof, triazole and substituted analogues thereof, pyrazole and substituted analogues thereof, imidazole compounds and substituted analogues thereof.

Another method of preparing natural and modified nucleosides made possible by the present invention is to use recombinant cells or corresponding crude or purified cell extracts to catalyse the phosphorolysis reaction of a donor nucleoside (using UdP or PNP, depending on the base present in the donor nucleoside) and obtain α-sugar-1-phosphate which can optionally be isolated by chromatography, extraction or precipitation techniques and used in the subsequent reaction of transferring the sugar onto a suitable acceptor base in the presence of UdP or PNP (depending on the nature of the acceptor base).

The availability of recombinant bacterial strains which overexpress the UdP and PNP enzymes separately also enables the conditions of the transglycosylation reactions to be fixed, in terms of optimum activity of each of the two enzymes, by means of preliminary tests in which the reaction is carried out in the presence of mixtures containing varying proportions of cells of each of the two strains. For each transglycosylation reaction it is therefore possible to define, on an analytical scale, the optimum ratios of UdP and PNP enzyme activity while, in the subsequent preparative scale-up, it is possible to use either a mixture of cells of the two strains that express UdP and PNP individually, or only the strain that co-expresses UdP and PNP if their ratios are already optimum, or optionally the strain that co-expresses UdP and PNP, integrated with cells of strains expressing UdP or PNP. Such optimisation of the reaction conditions can be carried out using crude or purified cell extracts prepared from the cell paste of recombinant strains overexpressing UdP and PNP.

By way of example of optimisation of the bioconversion reactions in the present invention, a detailed description is given of the procedures relating to the preparation of 9-β-D-arabinofuranosyladenine(Ara-A) and 1-β-D-ribofuranosyl-1,2,4-triazole-3-carboxamide (ribavirin) which indicated that the best results were obtained with UdP:PNP activity ratios of 2:1 and 1:1, respectively, and with a concentration of 10 units/ml of UdP and 5 units/ml of PNP for Ara-A and 10 units/ml of either UdP or PNP for ribavirin. These enzyme

activity ratios, or others found to be optimum for the reaction concerned, can be readily implemented using the recombinant strains described in the present invention, in order to optimise the concentration of cells to be used as biocatalysts, while at the same time obtaining the maximum bioconversion yield compatible with the constants of equilibrium of the enzyme reactions and a reduction in the reaction times. Analogously, it is possible to optimise all the transglycosylation reactions for the preparation of nucleosides and modified analogues thereof.

When the novel recombinant strains expressing the fusion proteins UdP-PNP or UdP-(L)-PNP (or the corresponding crude or purified extracts) are used for the bioconversion reactions, there is the advantage of using bifunctionals polipeptides in which the components having the activity of enzymes UdP and PNP are present in the stechiometric ratio 1:1. Furthermore, as nucleosides production via bioconversion is carried out by way of two successive reactions catalyzed respectively by UdP and PNP, the use of biocatalysts based on the bifunctionals fusion proteins UdP-PNP or UdP-(L)-PNP according to the present invention may improve the overall kinetic of the reactions thanks to a more efficient transfer of intermediates products from a reaction site to the other one.

The novel recombinant strains described in the present invention enable natural nucleosides and modified nucleosides to be prepared with significantly better results than those obtained by the enzyme techniques known hitherto which are based on the use of isolated enzymes or on the use of bacterial cells of wild-type micro-organism strains and cultivated micro-organism strains under conditions for inducing the activities of the phosphorylase enzymes.

A comparison of various transglycosylation reactions which were carried out using constant ratios between the concentration of donor nucleoside (60 mM) and acceptor base (20 mM) and in which a productivity parameter was calculated (Simon *et al.*, Angew.Chem 24, 539-553, 1985) which, in addition to specific activity, also takes into account operating factors, such as, for example, intra-cellular and extra-cellular transport phenomena and the volumetric concentration of the end products, indicates that the use of the recombinant strains or of the corresponding crude or purified extracts to which the present invention relates is always characterised by greater bioconversion efficiency and by higher productivity per unit of time and of volume compared with the use of conventional micro-organisms (Table 4).

Table 4.

Comparison of the efficiency of transglycosylation reactions catalysed by recombinant *Escherichia coli* cells (E) and by control *Enterobacter aerogenes* cells (C).

The reactions were carried out at 60°C for the time indicated, using the same concentrations of donor nucleoside (60 mM) and of acceptor base (20 mM). The bioconversion yield was calculated relative to the acceptor base by RP-HPLC analysis of the reaction mixture. The efficiency of the reaction is expressed by the productivity index P, calculated by the following formula $P = n \cdot m^{-1} \cdot t^{-1} \cdot 1000$ where n = concentration of the end product (g/l); m = wet cell paste (g/l of reaction mixture) and t = reaction time in hours.

Product	Nucleoside 60 mM	Base 20 mM	Cell paste g/100 ml		t hours		Bioconversion %		P	
			C	E	C	E	C	E	C	E
Ribavirin	Uridine	1,2,4-triazolo-3-carbox-amide	5	0.1	25	6	85	92	3	750
2'-deoxy-guanosine	2'-deoxy-uridine	Guanine	5	0.5	4	2	80	86	25	550
2'-deoxy-adenosine	2'-deoxy-uridine	Adenine	1	0.05	2	1	95	95	240	9600
Thymidine	2'-deoxy-uridine	Thymine	0.5	0.05	1	3	59	60	600	2000
2'-deoxy-ribofuranosyl-2,6-diamino-purine	2'-deoxy-uridine	2,6-diamino-purine	2	0.05	2	1.5	89	91	125	6660
Ara-A	Ara-U	Adenine	5	0.5	20	2	85	87	5	480

In particular, as shown in the example given in Table 5 regarding the preparation of Ara-A from Ara-U and adenine, the use of the recombinant strains enables conventional bioconversion processes to be improved both from the technical point of view and from the economic point of view and enables higher bioconversion yields, shorter reaction times, and a higher volumetric yield of end products to be obtained using a lower concentration of cells or corresponding crude or purified extract.

Table 5

Comparison of the operating conditions for the preparation of Ara-A by transglycosylation catalysed by recombinant *E.coli* cells and by a comparison *E.aerogenes* preparation.

Operating conditions	Recombinant *E.coli* Cells	*E. aerogenes* Cells
Strain	MG1655/pGM716 or DH5α/pGM716	Induced *E.aerogenes* ATCC 13048
Ara-U : Adenine ratio	75 :75 (mM)	40 : 40 (mM)
Cell concentration	0.5%	5%
Reaction time	4 hours	20 hours
Bioconversion yield	70%	55%
Volumetric yield	14 g Ara-A/litre	5 g Ara-A/litre

A further advantage derived from the use of the recombinant strains to which the present invention relates is the simplification of the processes for recovering and re-using the cell biomass or the corresponding crude or purified cell extract resulting from the presence of a lower cell concentration; thus, for example, any recovery of the cells or the extract by filtration or ultrafiltration and their subsequent recycling is considerably faster when the recombinant strains described in the present invention are used. In some cases, in particular when substrates having a high affinity for enzymes are used, the concentration of recombinant cells or of the corresponding crude or purified cell extract is reduced to such low values that it may be economically advantageous to avoid having to recover them, with a further simplification of the production process.

The purpose of the Examples given below is to illustrate the present invention without constituting a limitation of the field of application thereof.

Example No. 1

Cloning of the *udp* gene of *Escherichia coli* into an expression vector

The *E.coli udp* gene *sequence* was found in the EMBL data bank with the accession number X15689. The gene was amplified by PCR with the oligonucleotides 5'-ATCGGTACCATCCATGTCCAAGTCTGATGTTTTTCATCTC-3' and 5'-AGACGGTCGACAAGAGAATTACAGCAGACGACGC-3' from the *E.coli* strain K12 MG1655 (Singer *et al.*, Microbiol. Rev. 53, 1-24, 1989). The amplified region comprises the entire sequence of the *udp* gene starting from the start codon ATG up to 7 bp downstream of the stop codon TAA. A *Kpn*I restriction site was inserted at the 5' of the gene, followed by four bases selected at random. A *Sal*I site is present at the 3' of the gene. The amplified fragment, digested with *Kpn*I and *Sal*I, was cloned into the polylinker region of the pUC18 vector which carries the ampicillin resistance gene (Yanish and Perron, Gene 33, 103-119, 1985; EMBL accession number L08752). After transformation of the DH5α strain (Hanahan, J. Mol. Biol. 166, 557-580, 1983), the pGM679 plasmid was obtained (Figure 1). In the construct, a fusion is created between the first codons of the *lacZ* gene of pUC18 and the entire *udp* sequence (Figure 2) and the transcription is under the control of the *lac* promoter of the vector.

The cloned region was completely sequenced and it was found to be completely identical with the data bank sequence. The pGM679 plasmid sequence is listed.

The pBR322 Tet gene, which confers tetracycline resistance (Bolivar *et al.*, Gene 2, 95-113, 1977; EMBL accession number J01749) was then inserted into the pGM679 plasmid. The gene, preceded by its promoter, was obtained by *Hind*III digestion from the interposon pHP45W708-Tet (Fellay *et al.*, Gene 52, 147-154, 1987) and cloned into the *Hind*III site of pGM679. The resultant plasmid was named pGM708 (Figure 1). Its complete sequence is listed.

Example No. 2

Cloning of the *deoD* gene of *Escherichia coli* into an expression vector

The *E.coli deoD* gene sequence was found in the EMBL data bank with the accession number M60917. The gene was amplified by PCR with the oligonucleotides 5'-CTGAATTCTTCCATGGCTACCCCACACATTAATGCAG-3' and 5'-TCATGGTCGACTTACTCTTTATCGCCCAGCAGAACG-3' from the *E.coli* strain K12 MG1655 (Singer *et al.*, Microbiol. Rev. 53, 1-24, 1989). The amplified region comprises

the entire sequence of the *deoD* gene:starting from the start codon ATG up to the stop codon TAA. An *Eco*RI restriction site was inserted at the 5' of the gene, followed by four bases selected at random. A *Sal*I site is present at the 3' of the gene. The amplified fragment, digested with *Eco*RI and *Sal*I, was cloned into the polylinker region of the pUC18 vector, which carries the gene for ampicillin resistance (Yanish and Perron, Gene 33, 103-119, 1985; EMBL accession number L08752). After transformation of the DH5α strain (Hanahan, J. Mol. Biol. 166, 557-580, 1983), the pGM678 plasmid was obtained (Figure 1). In the construct, a fusion is created between the first codons of the *lacZ* gene of pUC18 and the entire *deoD* sequence (Figure 2) and the transcription is under the control of the *lac* promoter of the vector. The cloned region was completely sequenced and was found to be completely identical with the data bank sequence. The pGM678 plasmid sequence is listed.

The Tet gene, which confers tetracycline resistance, was then inserted into the pGM678 plasmid, in a manner analogous to that described in Example No. 1. The resultant plasmid was called pGM707 (Figure 1). Its complete sequence is listed.

The *deoD* gene was also cloned in a different vector as reported herebelow.

The region *Pvu*II-*Nde*I of pUC18 plasmid (end filled with Klenow) containing the replication origin was linked to the fragment *Eco*RI (filled)-*Hind*III (filled) containing the polylinker to obtain the resulting plasmid pGM746whose sequence is listed. pGM746 was subsequently digested with *Bam*HI (filled)-*Sph*I and linked to fragment *Nhe*I (filled)-*Sph*I of plasmid pGM709 in which is contained the *deoD* gene preceded by a Shine-Dalgarno sequence for the ribosome binding site (see example 3). The resulting plasmid was called pGM747 and its sequence is also listed.

The region containing the tac promoter was obtained by PCR amplification with oligonucleotides 5'-ATTGAGCTCGACATCATAACGGTTCTGGC and 5'-ATTGGATCCTGTGTGAAATTGTTATCCGC of plasmid pGZ119 (Lessl et al., J. Bacteriol. 174, 2493-2500, 1992), digestion of the fragment with *BamH*I-*Sac*I and insertion in *BamH*I-*Sac*I of pGM747 upstream *deoD*. The resulting plasmid pGM751 (figure 3) contains the *deoD* gene starting from *tac* promoter and expresses the PNP enzyme identical to the wild-type one. The pGM751 sequence is listed.

Example No. 3

Cloning of the *udp* and *deoD* genes into a single expression vector

The *udp* and *deoD* genes were cloned into the same vector in order to express the UdP and PNP enzymes simultaneously inside the same cell. This was effected by inserting the *deoD* gene into the pGM679 plasmid, downstream of *udp*. For the construction, the *EcoRI*-*SalI* fragment of pGM678, containing the *deoD* gene, was cloned into the pBAD24 vector (Guzman *et al.*, J. Bacteriol. 177, 4121-4230, 1995; EMBL accession number X81838) obtaining plasmid pGM709. The fragment *Nhe*I (with the ends filled) - *Sph*I of this construct was cloned into pGM679, digested *Sal*I (filled)-*Sph*I, to give pGM712 (Figure 1). In pGM712, both of the *udp* and *deoD* genes are transcribed starting from the *lac* promoter, but the translation of *deoD* is independent of that of *udp* because a sequence for the attachment of ribosomes is present upstream of *deoD* (Figure 2). It will be appreciated that the PNP protein expressed by pGM712 is identical to the wild protein because the fusion with the first codons of *lacZ* at the 5' of the gene was eliminated (Figure 2). The complete pGM712 sequence is listed.

The Tet gene, which confers tetracycline resistance, was subsequently inserted into the pGM712 plasmid as described in Example No. 1. The resultant plasmid was called pGM716 (Figure 1). Its complete sequence is listed.

The *udp* and *deoD* genes were also cloned in a different vector in which they are simultaneously expressed in this order starting from *tac* promoter, as herebelow reported.

The fragment *Sal*I-*Hind*III, obtained by PCR amplification using the pGM679 DNA as a template and the oligonucleotides 5'-TCCAGTCGACACAGGAAACAGCTATGA and 5'-TACGAAGCTTA AGAGAATTACAGCAGACG, was inserted into plasmid pGM751, digested with *Sal*I-*Hind*III, in order to obtain plasmid pGM800 bearing gene *udp* cloned downstream *deoD*. Both genes are transcribed starting from *ptac* but the transduction is independent. The complete sequence of pGM800 is listed.

The gene Tc for tetracycline resistance was subsequently inserted into pGM800 according to an analogous process to that reported in example 1, thus obtaining plasmid pGM807 (figure 3) whose sequence is also listed.

Example No. 4

Cloning of fusion proteins UdP-PNP and UdP-(L)-PNP

The sequence coding for UdP and PNP have been fused to each other either directly or separated by a short aminoacidic linker. The plasmids were obtained by subsequent steps starting from pGM716. In particular, plasmid pGM716 was digested with *Hpa*I and closed again so to have the deletion in the terminal part of gene *udp* and in the starting part of *deoD* and create plasmid pGM769 with a unique site *Hpa*I. The 3' portion of *udp* was amplified by PCR with the oligonucleotides 5'-GGCCGTTAACCGCACCCAGCAAGAG and 5'-AGCCATGGACAGCAGACGACGCGCC; the 5' portion of *deoD* was amplified in the same way with the oligonucleotides 5'-GCTGTCCATGGCTACCCCACACATTAAT and 5'-CCGGGTTAACTTTGGAATCGGTGCAGG. Subsequently, using the product of the two PCRs as a template and the two extreme sequences, the complete region was amplified: the obtained fragment creates a fusion between *udp* and *deoD*, replacing the *udp* stop codon with a codon for serine, followed by *deoD* ATG codon. The fragment was digested with *Hpa*I (site present at the two extremities) and cloned in pGM769 *Hpa*I site. The resulting plasmid was called pGM771 (figure 4). In pGM771, the fused protein UdP-PNP is then transcribed starting from *lac* promoter. The plasmid sequence is listed.

Plasmid pM771 was subsequently modified by inserting the 5'-CATGGGCGGT GGCAGCCCGGGCATTCTGGCCATG linker in the unique *Nco*I site, immediately upstream the starting *deoD* ATG. The resulting plasmid, called pGM795 (figure 4) expresses a fusion protein formed by UdP+ a 11 aminoacid linker (ser-met-gly-gly-gly-ser-pro-gly-ile-leu-ala) + PNP. The pGM795 sequence is listed.

Example No. 5

Transformation of *E.coli*

The *E.coli* strain K12 DH5α, which carries the *recA1* mutation (Hanahan, J.Mol.Biol. 166, 557-580, 1983) and the wild-type strain MG1655 (Singer *et al.*, Microbiol.Rev. 53, 1-24, 1989) were transformed with plasmids pGM678, pGM679, pGM707, pGM708, pGM712, pGM716, pGM771, pGM795, pGM751, pGM800 and pGM807. The genotype of the strains and some characteristics of the recombinant strains are given in Tables 6 and 7. The pGM678, pGM679, pGM712, pGM751 and pGM807 transformants were selected on medium containing ampicillin (50 μg /ml) and the pGM707, pGM708, pGM716,

pGM771,pGM795 and and pGM907. pGM771, pGM795 and pGM807 transformants were selected on medium containing tetracycline (12.5 μg /ml).

Table 6. Genotype of the host strains

Strain	Genotype	Reference
E.coli K12 DH5α	F',ϕ80dlacZΔM15, Δ(lacZYA-argF)U169, *deoR*, *recA1*, *endA1*, *hsdR17*(r_{K^-}, m_{K^+}), *phoA*, *supE44*, λ', *thi* 1, *gyrA96*, *relA1*	Hanahan, J. Mol. Biol. 166, 557-580, 1983
E.coli K12 MG1655	LAM-*rph-1*	Singer *et al.*, Microbiol. Rev. 53, 1-24, 1989

Table 7. Characteristics of the novel recombinant strains

Name of the strain	Expression of the cloned proteins	Resistence
DH5α/pGM678	purine nucleoside phosphorylase	ampicillin
DH5α/pGM679	uridine phosphorylase	ampicillin
DH5α/pGM707	purine nucleoside phosphorylase	tetracycline/ ampicillin
DH5α/pGM708	uridine phosphorylase	tetracycline/ampicillin
DH5α/pGM712	purine nucleoside phosphorylase and uridine phosphorylase	ampicillin
DH5α/pGM716	purine nucleoside phosphorylase and uridine phosphorylase	tetracycline/ampicillin
MG1655/pGM678	purine nucleoside phosphorylase	ampicillin
MG1655/pGM679	uridine phosphorylase	ampicillin
MG1655/pGM707	purine nucleoside phosphorylase	tetracycline/ampicillin
MG1655/pGM708	uridine phosphorylase	tetracycline/ampicillin
MG1655/pGM716	purine nucleoside phosphorylase and uridine phosphorylase	tetracycline/ampicillin
DH5α/pGM771	fusion protein UdP-PNP	tetracycline/ampicillin
DH5α/pGM795	fusion protein UdP-(L)-PNP	tetracycline/ampicillin
MG1655/pGM771	fusion protein UdP-PNP	tetracycline/ampicillin
MG1655/pGM795	fusion protein UdP-(L)-PNP	tetracycline/ampicillin
DH5α/pGM751	purina nucleoside phosphorylase	ampicillin
DH5α/pGM800	purine nucleoside phosphorylase and uridine phosphorylase	ampicillin
DH5α/pGM807	purine nucleoside phosphorylase and uridine phosphorylase	tetracycline/ampicillin

The presence of the plasmid in the transformed strains was confirmed by extraction of the plasmid DNA and analysis on 0.6% agarose gel.

The growth of the transformed strains in LD broth (composition per litre: 10 g Bactotryptone (Difco), 5 g Yeast extract (Difco), 5 g NaCl) or in solid medium (LD + 10 g/l agar), to which was added ampicillin (50 μg/ml) or tetracycline (12.5 μg/ml, only for the strains transformed with pGM707, pGM708, pGM716, pGM771, pGM795 and pGM807) is comparable to that of the control strains transformed with the pUC18 vector. In addition, the strains transformed with the plasmids pGM707, pGM708, pGM716, pGM771, pGM795 and pGM807, carryin both resistance genes, do not demonstrate differences in growth in the presence of ampicillin and tetracycline.

Example No. 6

Evaluation of the expression of the UdP and PNP proteins in the recombinant strains.

Precultures of the recombinant strains were obtained by inoculating single clones into LD medium to which an antibiotic had been added and by incubating without agitation at 37°C overnight. The cultures were diluted 1 : 20 in LD medium + antibiotic in a flat-bottomed flask and incubated at 37°C with agitation until the stationary phase was reached, corresponding to cell density values of approximately 2 units of optical density at 600 nm. The total proteins extracted from 1 ml of culture were separated on 15% polyacrylamide gel under reducing conditions (SDS-PAGE) and the proteins were visualised by staining with Coomassie Blue. The PNP and UdP proteins were identified on the basis of the molecular weight of approximately 26.6 kDa for PNP and 28.2 kDa for UdP. The result obtained from the extracts of strains MG1655/pGM707, pGM708 and pGM716 is given in Figure 5. Electrophoretic analysis shows that, in all the samples studied, overexpression of UdP and PNP has occurred, because the corresponding protein bands represent a significant percentage of the total cell proteins; this result is confirmed by the quantitative determination of the enzyme activities which is given in Tables 1 and 2 and by the quantitative determination of UdP and PNP expression effected by reverse phase high pressure liquid chromatography (RP-HPLC). For that purpose, the soluble extract was analysed on a C4-Vydac analytical column, dimensions 4.6 x 250 mm, using a mobile phase constituted by acetonitrile-H_2O containing 0.1% trifluoroacetic acid and operating in accordance with the following parameters: flow rate of 0.75 ml/minute; elution with a

gradient from 40% acetonitrile to 65% acetonitrile in 30 minutes; temperature of 45°C; UV detection at a wavelength of 215 nm. Under the analysis conditions applied, the elution times for UdP and PNP were approximately 13 minutes and 15 minutes, respectively. The quantitative determination was carried out by comparing the area of the peak of interest with the area of the peak of standard UdP and PNP preparations separated under the same conditions as the samples.

Because, in the recombinant strains, the *deoD* and *udp* genes are cloned under the control of the *lac* promoter, the growth of the cells and the expression of the UdP and PNP proteins were monitored both in the absence and in the presence of 40 mg/l of IPTG as transcription inducer. The results obtained indicated that the presence of IPTG does not modify cell growth and does not increase the level of PNP and UdP expression (possibly due to the insufficient amount of repressor in those strains). This last result indicates that, in the recombinant strains to which the present invention relates, the expression of the *deoD* and *udp* genes is constitutive and reaches very high levels without phenomena of cell damages or diminished cell vitality.

Example No. 7.

Determination of the enzyme activity of uridine phosphorylase and purine nucleoside phosphorylase expressed intracellularly in recombinant bacterial cells.

The strains were grown as described in Example No. 5. The cells were harvested by centrifugation, weighed in the form of wet cell paste and stored at -20°C until enzyme analysis was carried out.

The activity of the UdP enzyme was determined in a phosphorolysis test by incubating for 5 minutes at 30°C the soluble fraction (cell extract) obtained by sonication of a known amount of a suspension of the cell paste and by centrifugation of the homogenate in 100 mM-pH 7 phosphate buffer containing 60 mM of the uridine substrate. The enzyme reaction was blocked by acidification with 0.1N HCl; the suspension was filtered and analysed by RP-HPLC on a C18 column (Hypersyl 100; 4.6 x 250 mm), eluting under isocratic conditions with a mobile phase constituted by 0.02 M K_2HPO_4 in methanol-H_2O (4 : 96 v/v) and adjusted to pH 4.5 with NH_4OH. The amount of uracil formed in the reaction was determined by reference to a standard curve and the enzyme activity of the

cell preparation was calculated in μmol uracil/min/g wet cell paste (units/g). The activity of the PNP enzyme was determined in a phosphorolysis test by incubating for 10 minutes at 30°C the soluble fraction (cell extract) obtained by sonication of a known amount of a suspension of the cell paste and by centrifugation of the homogenate in 100 mM-pH 7 phosphate buffer containing 50 mM of the inosine substrate. The enzyme reaction was blocked by acidification with 0.1N HCl; the suspension was filtered and analysed by RP-HPLC on a C18 column (Hypersyl 100; 4.6 x 250 mm), eluting under isocratic conditions with a mobile phase constituted by 0.02 M K_2HPO_4 in methanol-H_2O (4 : 96 v/v) and adjusted to pH 4.5 with NH_4OH. The amount of hypoxanthine formed in the reaction was determined by reference to a standard curve and the enzyme activity of the cell preparation was calculated in μmol hypoxanthine/min/g wet cell paste (units/g).

Example No. 8

Fermentation of the recombinant strains.

The recombinant strains to which the present invention relates were cultivated at high biomass either under batch mode or under fed-batch mode fermentation conditions.

The batch-mode fermentations were carried out using a fermenter having a working volume of 10 litres which was filled with 9 litres of medium having the following composition (per litre): 0.6 g KH_2PO_4; 3.2 g K_2HPO_4; 20 g Soytone (Difco); 36 g yeast extract (Difco); 1 g $MgSO_4$-$7H_2O$; 0.0125 g tetracycline (or other antibiotic used as a selection marker) and which was inoculated with 1 litre of a bacterial suspension previously cultivated for 20 hours at 30°C in medium having the following composition, per litre: 20 g tryptone; 10 g yeast extract; 10 g NaCl; 0.0125 g tetracycline.

The fermentation was carried out in accordance with the following operating parameters: 30°C; air flow of 1 litre/litre of culture/minute; initial agitation 250 rev/min modified automatically to maintain a level of O_2 at 20% of the saturation concentration; pH maintained at 7 by additions of H_3PO_4 or NH_4OH; time 24 hours. When fermentation was complete, the culture medium was centrifuged, the cell pellet was washed in 30 mM-pH 7 phosphate buffer. The biomass obtained (40-50 grams of wet cell paste/litre of culture medium) was stored at -20°C until it was brought into use.

The fed-batch mode fermentations were carried out using a fermenter having a working volume of 10 litres which was filled with 7 litres of medium at pH 6.8-7 having the following composition, per litre: 13.3 g KH_2PO_4; 4 g $(NH_4)_2HPO_4$; 1.25 g Soytone (Difco); 0.125 g yeast extract (Difco); 1.7 g citric acid; 2.5 g glycerol; 1.5 g $MgSO_4$-$7H_2O$; 0.08 g $CaCl_2$; 0.01 g thiamine, 0.0125 g tetracycline (or other antibiotic selector); 0.08 g $FeSO_4$-$7H_2O$; 0.02 g $MnSO_4$-H_2O; 0.03 g $ZnSO_4$-$7H_2O$; 0.003 g H_3BO_3; 0.06 g $CuSO_4$-$5H_2O$; 0.008 g $CoCl_2$-$6H_2O$; 0.004 g $NaMoO_4$-$2H_2O$. The fermenter was inoculated with 1 litre of bacterial suspension previously cultivated for 18-20 hours at 30°C in medium having the following composition, per litre: 13.3 g KH_2PO_4; 4 g $(NH_4)2HPO_4$; 5 g Soytone (Difco); 1.7 g citric acid; 10 g glycerol; 0.01 g thiamine; 0.0125 g tetracycline; 0.05 g $CaCl_2$-$2H_2O$; 1 g $MgSO_4$-$7H_2O$; 0.03 g $FeSO_4$-$7H_2O$; 0.01 g $MnSO_4$-H_2O; 0.01 g $ZnSO_4$-$7H_2O$; 0.003 g H_3BO_3; 0.02 g $CuSO_4$-$5H_2O$; 0.002 g $CoCl_2$-$6H_2O$; 0.002 g $NaMoO_4$-$2H_2O$.

The fermentation was carried out in accordance with the following operating parameters: 30°C; air flow of 1-1.2 litre/litre of culture/minute; initial agitation 150 rev/min modified automatically to maintain a level of O_2 at 20% of the saturation concentration for approximately 8-10 hours (batch phase) and subsequently a level of O_2 at 10% of the saturation concentration (fed-batch phase); pH maintained at 6.8-7 by additions of H_3PO_4 or NH_4OH. During the fed-batch phase, the fermentation was automatically supplied with a total of 2 litres of a solution having the following composition, per litre: 400 g glycerol; 200 g Soytone; 20 g yeast extract; 3 g $MgSO_4$-$7H_2O$; 0.0125 g tetracycline. When fermentation was completed (after approximately 50 hours) the culture medium was centrifuged, the cell pellet was washed in 30 mM-pH 7 phosphate buffer. The biomass obtained (150-200 grams of wet cell paste/litre of culture medium) was stored at -20°C until it was brought into use.

Example No. 9

Transglycosylation reactions on a laboratory scale and calculation of the productivity index

The transglycosylation reactions were carried out using various sugar-donating nucleosides at a concentration of 60 mM (uridine, 2'-deoxyuridine, Ara-U) and various acceptor bases at a concentration of 20 mM (1,2,4-triazole-3-carboxamide, guanine, adenine, thymine, 2,6-diaminopurine) at pH 7 in phosphate buffer (30 mM) in the presence of various concentrations of cell paste or corresponding crude or purified extract derived either from

cultures of the control micro-organism *E.aerogenes* or from cultures of the recombinant *E.coli* strain MG1655/pGM716 which overexpresses the UdP and PNP enzymes. The reactions were carried out at 60°C for various periods of time (from 1 hour to 25 hours) and the percentage bioconversion, relative to the initial concentration of acceptor base, was determined by RP-HPLC analysis of the diluted reaction mixture. The results obtained are given in Table 2.

The productivity index P was calculated for each reaction by applying the following formula:

$$P = n \cdot m^{-1} \cdot t^{-1} \cdot 1000$$

where n = concentration of the end product (g/l)

m = wet cell paste (g/l of reaction mixture)

t = reaction time in hours.

The productivity index represents an overall measure of the efficiency of the reaction because it takes into account both the characteristics of the enzyme-substrate interaction itself and operating parameters, such as the reaction time, the amount of cells used and the volumetric yield of end product.

Example No. 10

Optimisation of the use of recombinant *E.coli* cells in transglycosylation reactions

The preparation of ribavirin starting from uridine (60 mM) and 1,2,4-triazole-3-carboxamide (40 mm) and of Ara-A starting from Ara-U (40 mM) and adenine (40 mM) were studied as examples of optimisation of the use of recombinant *E.coli* cells in bioconversion reactions. In each case, the reactions were carried out at 60°C in the presence of 30 mM of potassium phosphate at pH 7 and in the presence of various amounts of cell paste obtained by fermentation of the strains MG1655/pGM707 (overexpressing the UdP enzyme) and MG1655/pGM708 (overexpressing the PNP enzyme). At predetermined intervals, aliquots of the reaction mixture were taken and analysed by RP-HPLC in order to determine the percentage bioconversion (calculated relative to the concentration of acceptor base).

The study was initially carried out by incubating the reaction mixture for 20 hours in the presence of a limiting concentration of cell paste (with total enzyme activity equal to or less than 2 units/ml) and by operating in such a manner as to have ratios of UdP enzyme units and PNP enzyme units varying in the following proportions 5:1; 2:1; 1:1; 1:2; 1:5.

The results obtained in the two bioconversion reactions are given in Table 8.

Table 8. Study of the transglycosylation reaction conditions

The reactions were carried out for 20 hours at 60°C in the presence of limiting concentrations of cell paste.

Preparation of ribavirin			Preparation of Ara-A		
UdP	PNP	Bioconversion yield	UdP	PNP	Bioconversion yield
units/ml		%	units/ml		%
1	0.2	60.7	1	0.2	54.0
1	0.5	77.3	1	0.5	65.2
1	1	81.6	1	1	63.8
0.5	1	80.0	0.5	1	26.4
0.2	1	78.1	0.2	1	9.2

The results given in the Table demonstrate that the optimum UdP and PNP activity ratios are 1:1 and 1:0.5, respectively, for the reaction for the formation of ribavirin and Ara-A.

These data were confirmed in the subsequent study in which enzyme concentrations 10 times higher were used, with the same proportions being maintained between the UdP units and the PNP units; in this study, the reaction kinetics were also determined by taking samples of reaction mixture at intervals of 1 hour for RP-HPLC analysis and calculation of the percentage bioconversion.

Tables 9 and 10 show, for the ribavirin and Ara-A preparation reactions, respectively, the optimum parameters in terms of percentage bioconversion and reaction time for the various proportions of UdP and PNP studied.

Table 9. Optimisation of the reaction conditions for the preparation of ribavirin

UdP units/ml	PNP units/ml	Reaction time hours	Bioconversion %
10	2	20	89.4
10	5	4	89.5
10	10	2	91.2
5	10	2	91.2
2	10	2	91.1

Table 10. Optimisation of the reaction conditions for the preparation of Ara-A.

UdP units/ml	PNP units/ml	Reaction time hours	Bioconversion %
10	2	3	70.5
10	5	2	70.8
10	10	2	70.6
5	10	6	70.1
2	10	6	70.0

The results of the optimisation study indicate that ribavirin can be obtained in two hours with a bioconversion yield of 91% using 10 units/ml of either UdP or PNP while Ara-A can be obtained in two hours with a bioconversion yield of approximately 71% using 10 units/ml of UdP and 5 units/ml of PNP.

On the basis of the enzyme activity titre of the recombinant *E.coli* strains described in the present invention, it is therefore possible to calculate the amount of cell paste necessary to prepare ribavirin and Ara-A under optimum conditions. In the case, for example, of the strains MG1655/pGM707 and MG1655/pGM716 having the specific activities given in

Table I, 0.4 and 0.2 gram of wet cell paste/100 ml of reaction mixture, respectively, will be used for the preparation of ribavirin and Ara-A.

Example No. 11

Pilot-scale preparation of Ara-A by transglycosylation reaction carried out with the *E. aerogenes* comparison strain, with the recombinant *E.coli* strains and with the corresponding cell extracts.

The process for the preparation of Ara-A by transglycosylation catalysed by *E. aerogenes* cells or by recombinant cells of *E.coli* MG1655/pGM716 or DH5α/pGM716 overexpressing UdP and PNP was studied on a reaction scale of 1000 litres.

50 kg of wet cell paste obtained by fermenting *E. aerogenes* were resuspended in approximately 200 litres of 30 mM phosphate buffer at pH 7 and mixed with 800 litres of phosphate buffer in which had been dissolved at elevated temperature 5.4 kg of adenine (final concentration 40 mM) and 8.9 kg of Ara-U (final concentration 40 mM). The mixture was maintained at 60°C, with agitation, for 20 hours, diluted to approximately 3000 litres with hot H_2O and subjected to diafiltration on a membrane, collecting approximately 5000 litres of ultrafiltrate. The bioconversion yield determined by RP-HPLC was approximately 55%. The residue containing the cell paste is used for a subsequent reaction. The ultrafiltrate was concentrated to approximately 1000 litres and cooled to collect the precipitate constituted by Ara-A contaminated with non-reacted adenine (approximately 30 g of adenine per 100 g of Ara-A). 5 kg of Ara-A (total yield approximately 46%) with a degree of purity higher than 99.5% were finally obtained after crystallisation with H_2O.

5 kg of wet cell paste or the corresponding crude or purified extract obtained by fermenting the strain MG1655/pGM716 or the strain DH5α/pGM716 were resuspended in approximately 20 litres of 30 mM phosphate buffer at pH 7 and mixed with 980 litres of phosphate buffer in which had been dissolved at elevated temperature 10.1 kg of adenine (final concentration approximately 74.6 mM) and 18.3 kg of Ara-U (final concentration approximately 74.6 mM). The mixture was maintained at 60°C, with agitation, for 4 hours to obtain a bioconversion yield of approximately 70%. The cell paste was recovered in order to be used again in subsequent reactions by dilution at elevated temperature and diafiltration in accordance with the procedure described above. The ultrafiltrate was

concentrated to a volume of approximately 1000 litres, cooled to collect the precipitate constituted by Ara-A which was subsequently crystallised from water to obtain approximately 14 kg of Ara-A with a degree of purity higher than 99.5%. According to an alternative procedure, in which the cells were not recovered and the diafiltration step was omitted, at the end of the reaction the mixture was heated to approximately 90°C and filtered at elevated temperature to eliminate the cells, and the filtrate was cooled to precipitate Ara-A contaminated with non-reacted adenine (approximately 20 g of adenine per 100 g of Ara-A). 14 kg of Ara-A (total yield 65%) having a degree of purity higher than 99.5% were finally obtained after crystallisation from reaction of 1000 litres. Similar results were obtained starting from a mixture of the cell pastes or the corresponding crude or purified extracts obtained by fermenting the recombinant *E.coli* strains MG1655/p707 or MG1655/p708 and the strains DH5α/pGM707 or DH5α/pGM707 overexpressing UdP and PNP, respectively.